

This report presents implementations and analysis of three fundamental Markov Decision Process (MDP) problems

Markov Decision Process, Dynamic Programming, Value Iteration, Policy Iteration, Reinforcement Learning

Introduction Markov Decision Processes (MDPs) provide a mathematical framework for modeling decision-making

Dynamic Programming (DP) methods offer a powerful approach to solving MDPs when a complete model of the en

Value Iteration: Iteratively updates value estimates until convergence.

Policy Iteration: Alternates between policy evaluation and improvement.

In this lab, we investigate three problems:

Grid World: A 4×3 stochastic navigation problem with varying step costs.

Gbike Rental (Original): Inventory management with Poisson-distributed demands.

Gbike Rental (Modified): Extended version with operational constraints.

Methodology

Grid World Value Iteration The Grid World is a classic RL environment consisting of a 4×3 grid with:

Terminal States: +1 reward at $(0,3)$ and -1 reward at $(1,3)$

Wall: Impassable obstacle at $(1,1)$

State Action Transition Probabilities: With probability 0.8, take a random action in all directions; 0.1 up, 0.1 down, 0.1 left, 0.1 right.

